# Novelty, Trends and Questions

*John Graham-Cumming*

*Research Director, Sophos's Anti-Spam Task Force*

# Quick Overview

Latest news on content obfuscation

- How to spell V*I*A*G*R*A this week

Spam trends

- What can 10 months of spam tell us?

Tough questions for an anti-spam vendor

- Keeping sales people on their toes

# What I said last year

At LISA ’03…

“Every major email client to have adaptive filtering by end of 2004”

Microsoft Outlook lagging

“Spam will get simpler with less trickery”

Happening, but not as fast as I thought.

“Many spams will use CSS”

Yes, this has happened.

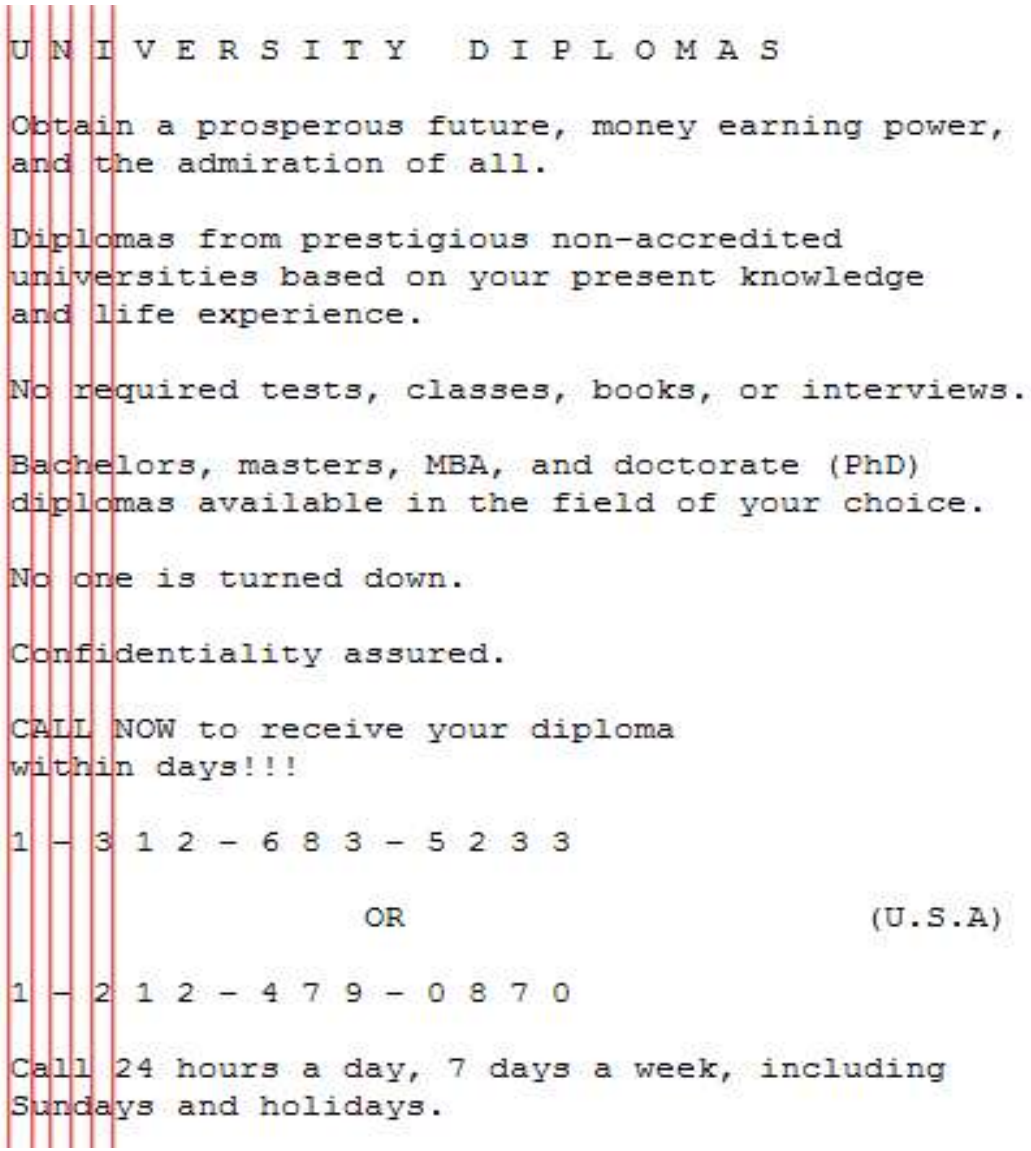
See Sophos whitepaper on CSS spam.

# Slice and Dice

Use <table> tag and monospace font to form text out of fragments

```
<table cellpadding=0 cellspacing=0 border=0><tr>
<td><table cellspacing=0 cellpadding=0 border=0><tr><td><font
    face="Courier New, Courier, mono" size=2>
 <br>U<br> <br>O<br>a<br> <br>D<br>u<br>a<br> 
    <br>N<br> <br>B<br>d<br> <br>N<br> <br>C<br>&nb
    sp;<br>C<br>w<br> <br>1<br> <br> <br> <br>
    1<br> <br>C<br>S<br></font></td></tr></table></td>
<td><table cellspacing=0 cellpadding=0 border=0><tr><td><font
    face="Courier New, Courier, mono" size=2>
   <br> N <br>   <br>bta<b
    r>nd <br>   <br>ipl<br>niv<br>nd <br>
       <br>o r<br>   <br>ach<br
    >ipl<br>   <br>o o<br>   <b
    r>onf<br>   <br>ALL<br>ith<br>   
    ;<br> -
     <br>   <br>   <br> &n
    bsp; <br> -
     <br>   <br>all<br>und<br></font></td></tr
    ></table></td>
<td><table cellspacing=0 cellpadding=0 border=0><tr><td><font
    face="Courier New, Courier, mono" size=2>
   <br>I V<br>   <br>in <b
    r>the<br>   <br>oma<br>ers<br>lif<br> &nbs
    p; <br>equ<br>   <br>elo<br>oma<br> &
    nbsp; <br>ne <br>   <br>ide<br> 
      <br> NO<br>in <br>   <br>
    3 1<br>   <br>   <br> 
      <br>2 1<br>   <br> 24<br>
    ays<br></font></td></tr></table></td>
<td><table cellspacing=0 cellpadding=0 border=0><tr><td><font
    face="Courier New, Courier, mono" size=2>
```

# Slice and Dice



U N I V E R S I T Y  D I P L O M A S

Obtain a prosperous future, money earning power,
and the admiration of all.

Diplomas from prestigious non-accredited
universities based on your present knowledge
and life experience.

No required tests, classes, books, or interviews.

Bachelors, masters, MBA, and doctorate (PhD)
diplomas available in the field of your choice.

No one is turned down.

Confidentiality assured.

CALL NOW to receive your diploma
within days!!!

1 - 3 1 2 - 6 8 3 - 5 2 3 3

OR (U.S.A)

1 - 2 1 2 - 4 7 9 - 0 8 7 0

Call 24 hours a day, 7 days a week, including
Sundays and holidays.

# Latest Trickery

# Staying Up To Date

New: Flex Hex, Sound of Silence, Blankety Blank

Stay up to date on spammer tricks…

…Read The Spammers' Compendium

www.jgc.org/tsc/

…Sign up for the free Anti-Spam Newsletter

www.jgc.org/

# Recap: Invisible Ink

Use HTML font colors to write white on white

```
<body bgcolor=white>
Viagra
<font color=white>Hi, Johnny!  It was
really nice to have dinner with you
last night. See you soon, love
Mom</font>
</body>
```

# Flex Hex: Exploiting an IE Bug/Feature

Recently spammers have been refreshing old tricks like “Invisible Ink” and “Camouflage” by exploiting Internet Explorer’s color handling.

Pad RHS to multiple of 3 characters; then divide into three parts

`#0F0` is the same as `#000F00`

`#F` is the same as `#0F0000`

`#0F0F` is the same as `#0F0F00`

But that’s the tip of the iceberg…

# Flex Hex

Pretend all non-hexadecimal characters are 0

`#zFtygn` is the same as `#0F0000`

`#zFt` is the same as `#000F00`

With more than 6 characters apply the multiple of three rule; truncate to a DWORD; take the MSB

`#6db6ec49efd278cd0bc92d1e5e072d68` is the same as `#6ecde0`

Lots of opportunity to fool a spam filter

Further discussion at `http://scrappy-do.blogspot.com`

# Sound of Silence

A type of web bug that uses the BGSOUND tag to track when an email is opened.

BGSOUND is used to add a background sound to a web page.

Typically used by spammers with a “silent” sound.

```
<BGSOUND SRC=“http://spammer-
  site.com/my@email.address”>
```

# Blankety Blank

- Using non-existent zero-width images to break up words
- A classic “hide bad words” technique
- Zero-width means that image is not visible to the user and no error is displayed for the missing image

```
Via<img src=“elephant.gif” width=“0”
  height=“2” border=“0”>gra
```

# Trends

# Ten Months of Spam

Sophos gathers spam every day through a network of “honeypots”

For this presentation crunched data from July 2003 through April 2004

- 208 Gb of spam
- 28 million individual spam messages

Examine the trends in spammer content trickery

Post-April 2004

- Spammers continue to innovate
- But many trends remain same

# Types of Content Trickery

Try to...

- Hide “bad” words so that filters ignore them
- Include many “good” words to distract the filter
- Obscure web site addresses to prevent recognition (10% of spams)

U*s*e  S I M P L E  trickéry

- 20% of spams

Write messages using HTML

# HTML Comments

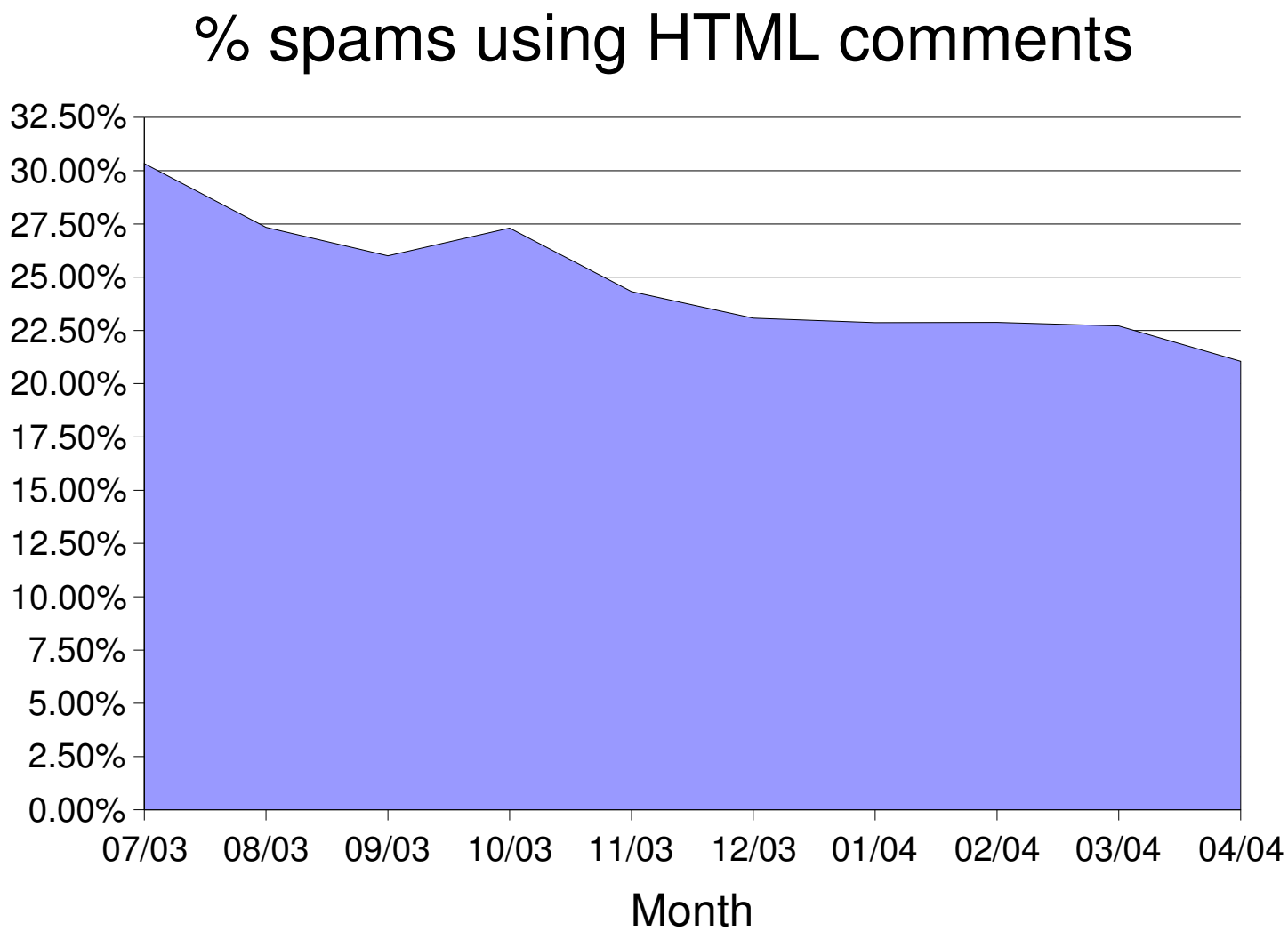
Uses HTML's commenting mechanism to break up bad words

HTML comments are written **`<!-- comment -->`** and the entire sequence is ignored and not displayed.

Easy to break up a word like Viagra:

```
V<!-- comment -->i<!-- comment --
>a<!-- comment -->g<!-- comment --
>r<!-- comment -->a
```

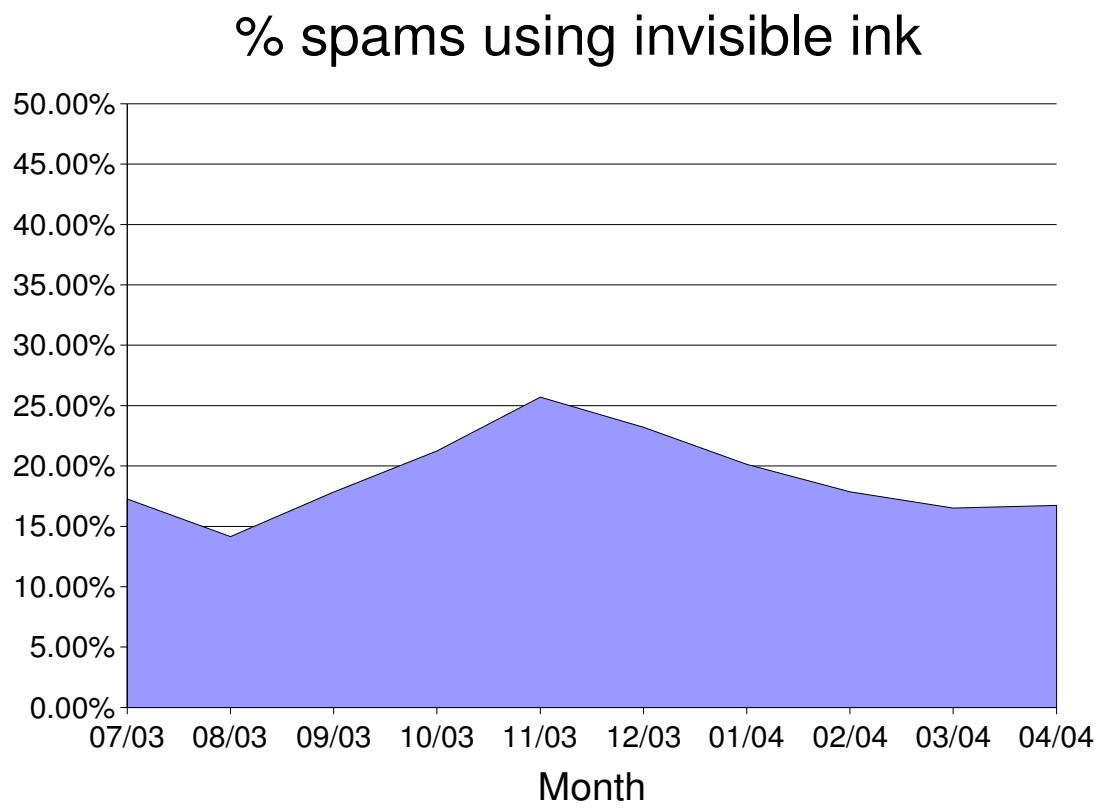
# HTML Comments Declining



# Invisible Ink

Remains a spammer favourite

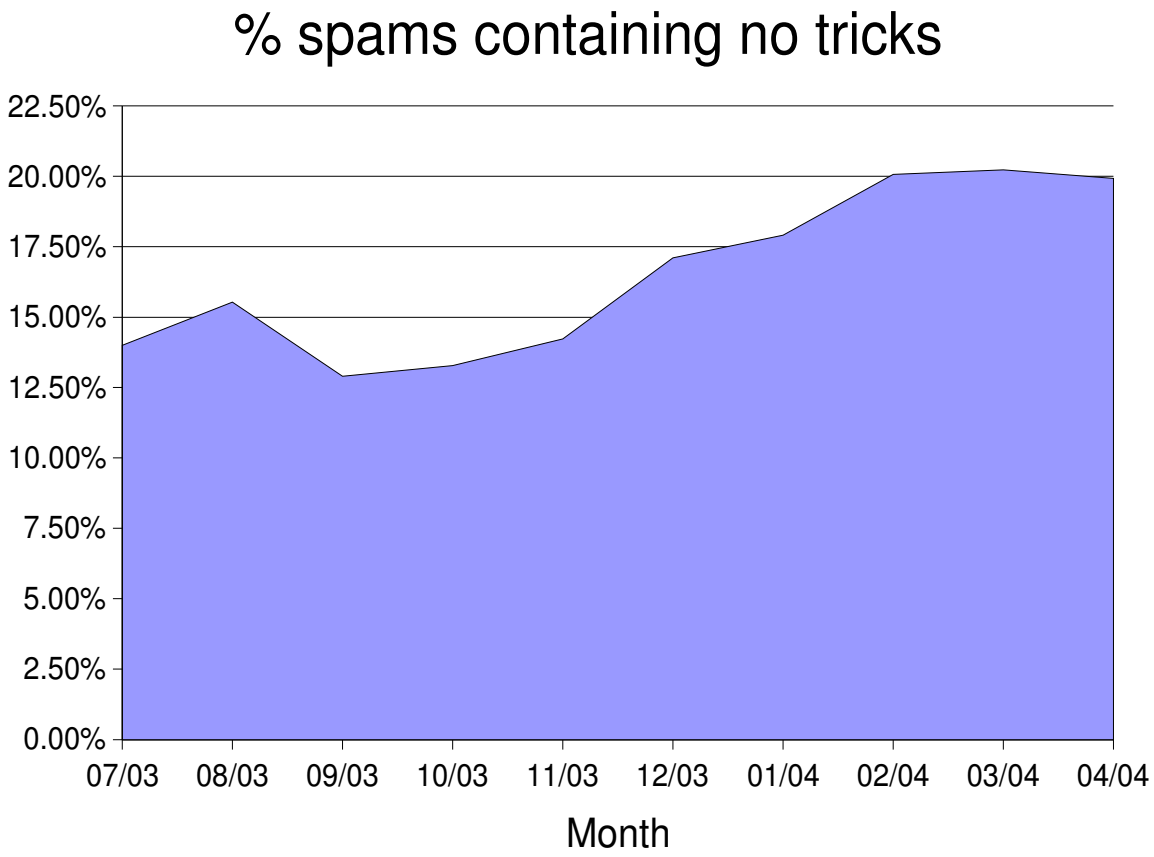
Used to insert hidden “good” words



# How Many Tricks?

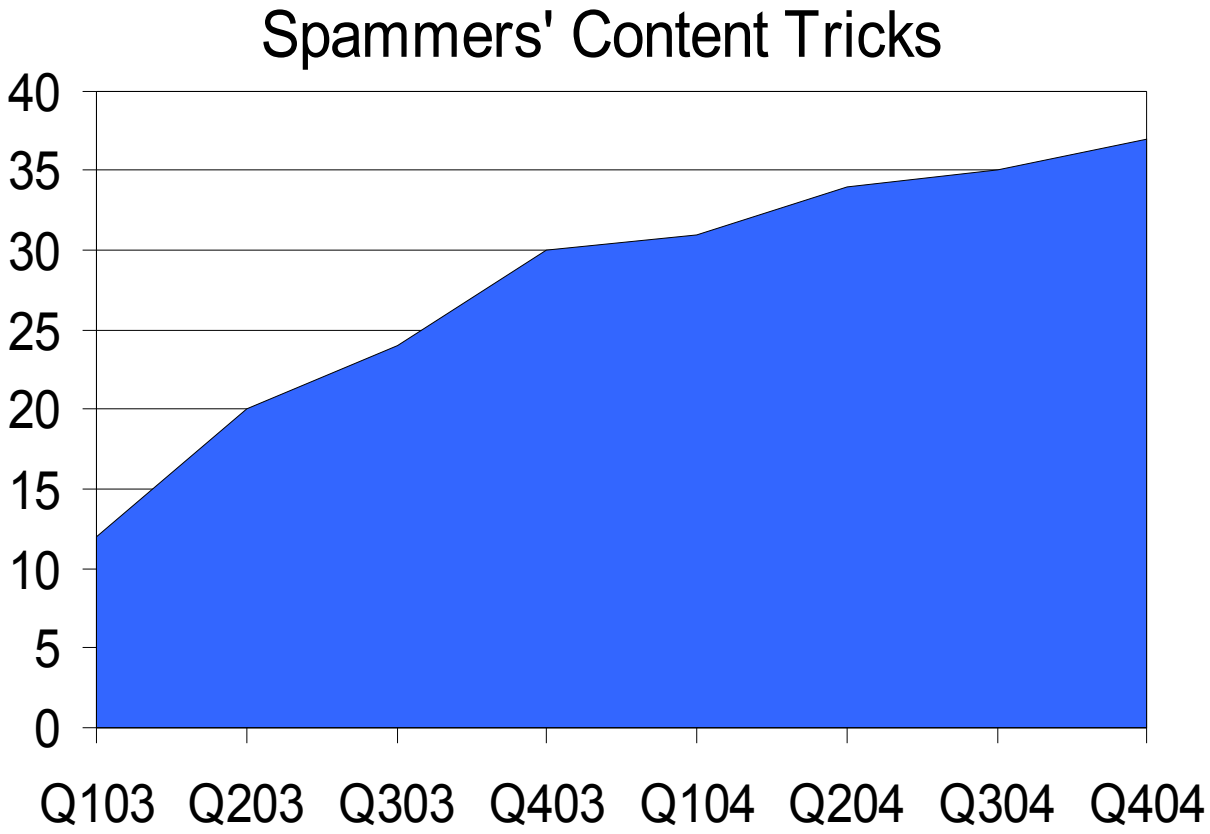
80% of spam incorporates trickery or obfuscation

But trick-free spams are increasing



# Trick Innovation

Spammer content-trickery is tracked in The Spammers' Compendium: www.jgc.org/tsc/



# More Trends

Further analysis of these trends available

My Virus Bulletin 2004 paper and presentation

Both available on www.jgc.org/

# Seven Tough Questions

# Question 1

How do you measure your system's **false positive rate**?

- False positive = good mail filtered as spam
- A key measure of a spam filter's effectiveness
- People are very sensitive to false positives ("I lost legitimate mail because of a spam filter")

Suppose spam filter vendor claims 99.999% accuracy

- Means 1 in 100000 wrongly filtered
- How do you verify that number?
- How much good mail is being filtered?

# Question 2

How often do you **react to changes** in spammer techniques?

Spammers are innovating constantly

- Different obfuscation tricks monthly
- Each spam is individually tailored
- Sending from zombie networks
- Moving web hosting daily

Spam filter must update at least daily, better if there are intra-day updates

# Question 3

What are your **top two ways** of catching spam?

Question designed to cut through the hype

- Each spam filter has specific strengths and heritage
- Spam filters that rely on identifying “known spam strings” tend to be weak and easily circumvented
- Challenge/response systems painful for the sender
- Most effective systems
  - sender blacklists
  - destination URL filtering
  - distributed checksumming
  - adaptive (Bayesian) filtering

# Question 4

Does you software **prevent web bugs** from firing?

No feedback to spammers

Don’t allow spammers to know if you saw the spam

- Spammers using BGSOUND, IMG, IFRAME, FRAME

Quarantined spam must be sanitized

# Question 5

How do you handle **legitimate bulk** mailings?

Just because a mail is bulk, it may not be spam

- Newsletters
- Mailing lists

Does the system handle this automatically?

Is it configurable by the sysadmin?

# Question 6

What happens if someone reports a **legitimate mail** as spam?

Common for people to “unsubscribe” by reporting legitimate mail as spam to make it go away

If I “unsubscribe” like this, does if affect other users

# Question 7

How does your system handle **non-English** spam and ham messages?

Simplistic filters may delete all non-English mail

Global organizations need English and non-English treated equally

# Conclusion

# Conclusion

Spammers are constantly adjusting their obfuscations

- They are innovating
- They are responding as filters improve
- They are not stupid!

Some spammers have realized that trickery makes their spams easier to spot

# Future

- Future spam
  - 2005: Savvy spammers reduce use of content obfuscations
- In Summary
  - Expect continued spammer innovation
  - Expect the set of tricks used by spammers to be a moving target as some fall out of favor and others are added